

Using Statistics in Ratio Study Analysis

David Schwab
Senior Statistician
Department of Local Government Finance
September 2010



- Mann-Whitney Test
- Spearman Rank Test
- Correlation
- Statistical Significance



Mann-Whitney Test

- Introductory Problem
- Concept
- Calculation
- Worked Example



Introductory Problem

- A taxpayer claims that properties in his neighborhood have been assessed inequitably. In particular, he alleges that sold properties were reassessed to bring them closer to the sale price, while unsold properties were left alone. To prove his case, he calculates the percentage change in total assessed value from last year to this year for both sold and unsold properties in his neighborhood.
- How can we determine whether his allegations are legitimate?



Introductory Problem

Sold Parcels	<u>Unsold</u>	<u>Parcels</u>	
4.9%	5.0%	5.0%	
4.6%	5.1%	5.3%	
4.8%	5.5%	4.9%	
5.0%	5.1%	4.8%	
12.3%	4.7%	5.4%	
13.0%	5.3%	4.5%	
24.5%	4.6%	4.6%	
	5.4%	5.3%	
	4.6%	5.2%	Median (Sold): 5.0%
	5.4%	4.8%	Median (Unsold): 5.1%
	5.1%	8.4%	
	5.5%	9.5%	
	4.8%	11.7%	
	5.1%	13.6%	
	4.8%	25.6%	



Sold Parcels	<u>Unsold</u>	<u>Parcels</u>	
10.0%	0.0%	0.0%	
10.0%	0.0%	0.0%	
10.0%	0.0%	0.0%	
10.0%	0.0%	0.0%	
10.0%	0.0%	0.0%	
10.0%	0.0%	0.0%	
10.0%	0.0%	0.0%	
	0.0%	0.0%	
	0.0%	0.0%	
	0.0%	0.0%	Median (Sold): 10.0%
	0.0%	0.0%	Median (Unsold): 0.0%
	0.0%	0.0%	
	0.0%	0.0%	
	0.0%	0.0%	
	0.0%	0.0%	



Sold Parcels	<u>Unsold</u>	<u>Parcels</u>	
10.0%	3.0%	3.0%	
10.0%	3.0%	3.0%	
10.0%	3.0%	3.0%	
10.0%	3.0%	3.0%	
10.0%	3.0%	3.0%	
10.0%	3.0%	3.0%	
10.0%	3.0%	3.0%	
	3.0%	3.0%	
	3.0%	3.0%	
	3.0%	3.0%	
	3.0%	3.0%	Median (Sold): 10.0%
	3.0%	3.0%	Median (Unsold): 3.0%
	3.0%	3.0%	
	3.0%	3.0%	
	3.0%	3.0%	



	<u>Parcels</u>	<u>Unsold</u>	Sold Parcels
	4.0%	2.6%	9.6%
	3.6%	3.1%	10.1%
	4.0%	3.5%	10.5%
	3.4%	3.4%	10.4%
	2.9%	3.7%	10.7%
	2.7%	3.4%	10.4%
	3.2%	2.9%	9.9%
	3.8%	3.5%	
Madian (Sold): 10 49/	3.4%	3.0%	
Median (Sold): 10.4% Median (Unsold): 3.4%	3.4%	3.1%	
Wicaidii (Olisola). 3.470	3.1%	2.8%	
	3.3%	3.6%	
	3.1%	2.8%	
	3.2%	3.6%	
	2.8%	3.8%	



- Percentage
 Change in assessed value for sold parcels (Sold).
- Percentage
 Change in
 assessed value for
 unsold parcels
 (Unsold).

Sold	Unsold
12%	5%
3%	1%
7%	0%
-7%	6%
4%	-2%



 Arrange all values in a single row, from largest to smallest.

12 7 6 5 4 3 1 0 -2 -7

Sold	Unsold
12%	5%
3%	1%
7%	0%
-7%	6%
4%	-2%



 Place an "S" under each Sold Value

Sold	Unsold
12%	5%
3%	1%
7%	0%
-7%	6%
4%	-2%



 Place a "U" under each Unsold value.

12 7 6 5 4 3 1 0 -2 -7 S S U U S S U U U S

Sold	Unsold
12%	5%
3%	1%
7%	0%
-7%	6%
4%	-2%



Count how many
 "U"s are to the right
 of each "S" (vice
 versa)

12 7 6 5 4 3 1 0 -2 -7 S S U U S S U U U S

Sold	Unsold
12%	5%
3%	1%
7%	0%
-7%	6%
4%	-2%

Rank S: 16

Rank U: 9

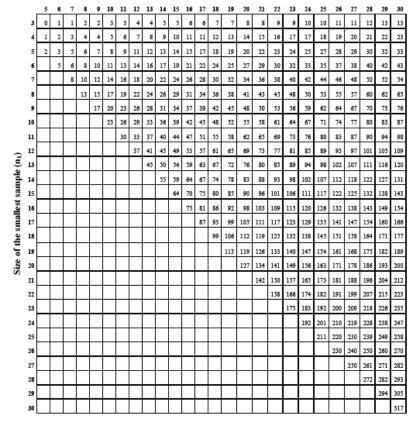


- Compare the value on the Mann-Whitney table with the smaller of the two ranks.
- If the rank is less than the lookup value, the sold and unsold may have been assessed differently.

Critical Values for the Mann-Whitney U-Test

Level of significance: 5% (P = 0.05)

Size of the largest sample (n₂)



The Open Door Web Site

Size of the largest sample (n₂)

		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	3	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	66	88	9	9	10	10	11	11	12	13	13
	4	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14	15	16	17	17	18	19	20	21	22	23
	5	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20	22	23	24	25	27	28	29	30	32	33
	6		5	6	8	10	11	13	14	16	17	19	21	22	24	25	27	29	30	32	33	35	37	38	40	42	43
	7			8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54
	8				13	15	17	19	22	24	26	29	31	34	36	38	41	43	45	48	50	53	55	57	60	62	63
	9					17	20	23	26	28	31	34	37	39	42	45	48	50	53	56	59	62	64	67	70	73	76
	10						23	26	29	33	36	39	42	45	48	52	55	58	61	64	67	71	74	77	80	83	87
	11							30	33	37	40	44	47	51	55	58	62	65	69	73	76	80	83	87	90	94	98
~	12								37	41	45	49	53	57	61	65	69	73	77	81	85	89	93	97	101	105	109
of the smallest sample (n ₁)	13									45	50	54	59	63	67	72	76	80	85	89	94	98	102	107	111	116	120
ď	14										55	59	64	67	74	78	83	88	93	98	102	107	112	118	122	127	131
Sar	15											64	70	75	80	85	90	96	101	106	111	117	122	125	132	138	143
llest	16												75	81	86	92	98	103	109	115	120	126	132	138	143	149	154
ma	17													87	93	99	105	111	117	123	129	135	141	147	154	160	166
hes	18														99	106	112	119	125	132	138	145	151	158	164	171	177
oft	19															113	119	126	133	140	147	154	161	168	175	182	189
Size	20																127	134	141	149	156	163	171	178	186	193	200
Ø	21																	142	150	157	165	173	181	188	196	204	212
	22																		158	166	174	182	191	199	207	215	223
	23																			175	183	192	200	209	218	226	235
	24																				192	201	210	219	228	238	247
	25																					211	220	230	239	249	258
	26																						230	240	250	260	270
	27																							250	261	271	282
	28																								272	282	293
	29																									294	305
	30																										317



- A taxpayer claims that properties in his neighborhood have been assessed inequitably. In particular, he alleges that sold properties were reassessed to bring them closer to the sale price, while unsold properties were left alone. To prove his case, he calculates the percentage change in total assessed value from last year to this year for both sold and unsold properties in his neighborhood.
- How can we determine whether his allegations are legitimate?

16



Sold Parcels	<u>Unsold</u>	<u>Parcels</u>	
4.9%	5.0%	5.0%	
4.6%	5.1%	5.3%	
4.8%	5.5%	4.9%	
5.0%	5.1%	4.8%	
12.3%	4.7%	5.4%	
13.0%	5.3%	4.5%	
24.5%	4.6%	4.6%	
	5.4%	5.3%	
	4.6%	5.2%	Median (Sold): 5.0%
	5.4%	4.8%	Median (Unsold): 5.1%
	5.1%	8.4%	. ,
	5.5%	9.5%	
	4.8%	11.7%	
	5.1%	13.6%	
	4.8%	25.6%	



25.6	24.5	13.8	13.0	12.3	11.7	9.5	8.4	5.5	
5.5	5.4	5.4	5.4	5.3	5.3	5.3	5.2	5.1	
5.1	5.1	5.1	5.0	5.0	5.0	4.9	4.9	4.8	
4.8	4.8	4.8	4.8	4.7	4.6	4.6	4.6	4.6	4.5



25.6	24.5	13.8	13.0	12.3	11.7	9.5	8.4	5.5	
U	S	U	S	S	U	U	U	U	
5.5	5.4	5.4	5.4	5.3	5.3	5.3	5.2	5.1	
U	U	U	U	U	U	U	U	U	
5.1	5.1	5.1	5.0	5.0	5.0	4.9	4.9	4.8	
U	U	U	U	U	S	U	S	U	
4.8	4.8	4.8	4.8	4.7	4.6	4.6	4.6	4.6	4.5
U	U	S	U	U	U	S	U	U	U



25.6	24.5	13.8	13.0	12.3	11.7	9.5	8.4	5.5	
U	S	U	S	S	U	U	U	U	
5.5	5.4	5.4	5.4	5.3	5.3	5.3	5.2	5.1	
U	U	U	U	U	U	U	U	U	
5.1	5.1	5.1	5.0	5.0	5.0	4.9	4.9	4.8	
U	U	U	U	U	S	U	S	U	
4.8	4.8	4.8	4.8	4.7	4.6	4.6	4.6	4.6	4.5
U	U	S	U	U	U	S	U	U	U

Rank S: 117

Rank U: 93

Size of the largest sample (n₂)

		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	3	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	66	88	9	9	10	10	11	11	12	13	13
	4	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14	15	16	17	17	18	19	20	21	22	23
	5	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20	22	23	24	25	27	28	29	30	32	33
	6		5	6	8	10	11	13	14	16	17	19	21	22	24	25	27	29	30	32	33	35	37	38	40	42	43
	7			8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54
	8				13	15	17	19	22	24	26	29	31	34	36	38	41	43	45	48	50	53	55	57	60	62	63
	9					17	20	23	26	28	31	34	37	39	42	45	48	50	53	56	59	62	64	67	70	73	76
	10						23	26	29	33	36	39	42	45	48	52	55	58	61	64	67	71	74	77	80	83	87
	11							30	33	37	40	44	47	51	55	58	62	65	69	73	76	80	83	87	90	94	98
~	12								37	41	45	49	53	57	61	65	69	73	77	81	85	89	93	97	101	105	109
of the smallest sample (n ₁)	13									45	50	54	59	63	67	72	76	80	85	89	94	98	102	107	111	116	120
ď	14										55	59	64	67	74	78	83	88	93	98	102	107	112	118	122	127	131
Sar	15											64	70	75	80	85	90	96	101	106	111	117	122	125	132	138	143
llest	16												75	81	86	92	98	103	109	115	120	126	132	138	143	149	154
ma	17													87	93	99	105	111	117	123	129	135	141	147	154	160	166
hes	18														99	106	112	119	125	132	138	145	151	158	164	171	177
oft	19															113	119	126	133	140	147	154	161	168	175	182	189
Size	20																127	134	141	149	156	163	171	178	186	193	200
Ø	21																	142	150	157	165	173	181	188	196	204	212
	22																		158	166	174	182	191	199	207	215	223
	23																			175	183	192	200	209	218	226	235
	24																				192	201	210	219	228	238	247
	25																					211	220	230	239	249	258
	26																						230	240	250	260	270
	27																							250	261	271	282
	28																								272	282	293
	29																									294	305
	30																										317



Questions?



Spearman Rank Test

- Introductory Problem
- Concept
- Calculation
- Worked Example



Introductory Problem

- You have assessed a group of six properties, and all statistics are fine except for the PRD, which is low. You have re-checked your data and believe the assessments are correct.
- How can you be sure there is no inequity?



Introductory Problem

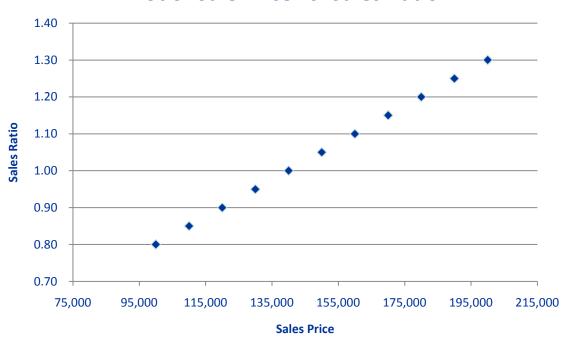
Assessed Value	Sale Price	<u>Ratio</u>	
35,000	56,000	0.63	Median: 1.04 COD: 16.5% PRD: 0.96
34,100	33,200	1.03	
91,000	86,400	1.05	
119,800	129,800	0.92	
136,600	101,200	1.35	
174,500	145,300	1.20	



- Vertical Inequity: When high-value properties are assessed by a different standard than lowvalue properties.
 - Progressivity: High-value properties overassessed
 - Regressivity: Low-value properties overassessed



Plot of Sale Price vs. Sales Ratio



<u>Ratio</u>
0.80
0.85
0.90
0.95
1.00
1.05
1.10
1.15
1.20
1.25
1.30

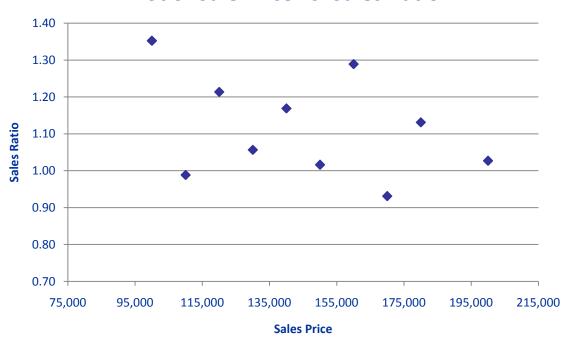




Sale Price	<u>Ratio</u>
100,000	1.30
110,000	1.25
120,000	1.20
130,000	1.15
140,000	1.10
150,000	1.05
160,000	1.00
170,000	0.95
180,000	0.90
190,000	0.85
200,000	0.80

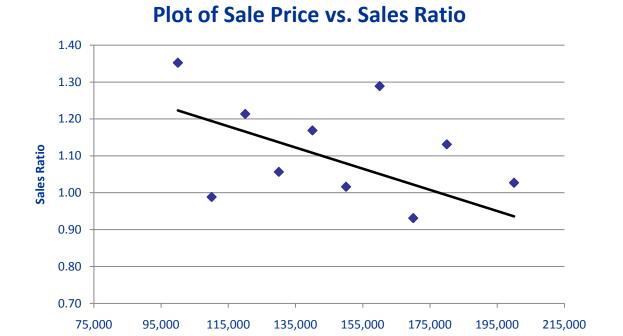


Plot of Sale Price vs. Sales Ratio



Sale Price	<u>Ratio</u>
100,000	1.35
110,000	0.99
120,000	1.21
130,000	1.06
140,000	1.17
150,000	1.02
160,000	1.29
170,000	0.93
180,000	1.13
190,000	0.70
200,000	1.03





Sales Price

Sale Price	<u>Ratio</u>
100,000	1.35
110,000	0.99
120,000	1.21
130,000	1.06
140,000	1.17
150,000	1.02
160,000	1.29
170,000	0.93
180,000	1.13
190,000	0.70
200,000	1.03



Rank the Sales Prices from highest (1) to lowest (11).

<u>Rank</u>	Sale Price	Ratio
11	100,000	1.35
10	110,000	0.99
9	120,000	1.21
8	130,000	1.06
7	140,000	1.17
6	150,000	1.02
5	160,000	1.29
4	170,000	0.93
3	180,000	1.13
2	190,000	0.70
1	200,000	1.03
	11 10 9 8 7 6 5 4 3	11 100,000 10 110,000 9 120,000 8 130,000 7 140,000 6 150,000 5 160,000 4 170,000 3 180,000 2 190,000



Next, rank the Ratios from highest (1) to lowest (11).

Sale Price	Rank	<u>Ratio</u>	<u>Rank</u>	Sale Price	Ratio
100,000	11	1.35	1	100,000	1.35
110,000	10	0.99	9	110,000	0.99
120,000	9	1.21	3	120,000	1.21
130,000	8	1.06	6	130,000	1.06
140,000	7	1.17	4	140,000	1.17
150,000	6	1.02	8	150,000	1.02
160,000	5	1.29	2	160,000	1.29
170,000	4	0.93	10	170,000	0.93
180,000	3	1.13	5	180,000	1.13
190,000	2	0.70	11	190,000	0.70
200,000	1	1.03	7	200,000	1.03



 Net, subtract each Ratio Rank from its corresponding Sales Price Rank.

Sale Price	<u>Rank</u>	Ratio	<u>Rank</u>	<u>Diff.</u>	Sale Price	Ratio
100,000	11	1.35	1	10	100,000	1.35
110,000	10	0.99	9	1	110,000	0.99
120,000	9	1.21	3	6	120,000	1.21
130,000	8	1.06	6	2	130,000	1.06
140,000	7	1.17	4	3	140,000	1.17
150,000	6	1.02	8	2	150,000	1.02
160,000	5	1.29	2	3	160,000	1.29
170,000	4	0.93	10	-6	170,000	0.93
180,000	3	1.13	5	-2	180,000	1.13
190,000	2	0.70	11	-8	190,000	0.70
200,000	1	1.03	7	-6	200,000	1.03



Square each of these differences.

					<u>Diff</u>		
Sale Price	<u>Rank</u>	<u>Ratio</u>	<u>Rank</u>	Diff.	(Squared)	Sale Price	<u>Ratio</u>
100,000	11	1.35	1	10	100	100,000	1.35
110,000	10	0.99	9	1	1	110,000	0.99
120,000	9	1.21	3	6	36	120,000	1.21
130,000	8	1.06	6	2	4	130,000	1.06
140,000	7	1.17	4	3	9	140,000	1.17
150,000	6	1.02	8	2	4	150,000	1.02
160,000	5	1.29	2	3	9	160,000	1.29
170,000	4	0.93	10	-6	36	170,000	0.93
180,000	3	1.13	5	-2	4	180,000	1.13
190,000	2	0.70	11	-8	64	190,000	0.70
200,000	1	1.03	7	-6	36	200,000	1.03



Add them all up.

					<u>Diff</u>		
Sale Price	Rank	<u>Ratio</u>	Rank	Diff.	(Squared)	Sale Price	<u>Ratio</u>
100,000	11	1.35	1	10	100	100,000	1.35
110,000	10	0.99	9	1	1	110,000	0.99
120,000	9	1.21	3	6	36	120,000	1.21
130,000	8	1.06	6	2	4	130,000	1.06
140,000	7	1.17	4	3	9	140,000	1.17
150,000	6	1.02	8	2	4	150,000	1.02
160,000	5	1.29	2	3	9	160,000	1.29
170,000	4	0.93	10	-6	36	170,000	0.93
180,000	3	1.13	5	-2	4	180,000	1.13
190,000	2	0.70	11	-8	64	190,000	0.70
200,000	1	1.03	7	-6	36	200,000	1.03

303



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6\frac{d}{n(n^2 - 1)}$$

d: 303

n: 11



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6\frac{303}{n(n^2 - 1)}$$



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6 \frac{303}{11(11^2 - 1)}$$



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6\frac{303}{1320}$$

$$11(11^2-1)$$

$$11(121-1)$$



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6(0.2295)$$

$$r = 1 - 1.3773$$

$$r = -0.3773$$



Now, plug the values for "r" and "n" into this formula.

 $t = r\sqrt{\frac{n-2}{1-r^2}}$

d: 303 n: 11 r: -0.3773



Now, plug the values for "r" and "n" into this formula.

d: 303n: 11r: -0.3773

$$t = -0.3773 \sqrt{\frac{11-2}{1-(-0.3773)^2}}$$



Now, plug the values for "r" and "n" into this formula.

d: 303n: 11r: -0.3773

$$t = -0.3773\sqrt{\frac{8}{1 - (-0.3773)^2}}$$



Now, plug the values for "r" and "n" into this formula.

$$t = -0.3773\sqrt{\frac{8}{0.8576}}$$

d: 303n: 11

r: -0.3773

$$1 - (-0.3773)^2$$

$$1 - 0.1424$$

0.8576



Now, plug the values for "r" and "n" into this formula.

$$t = -0.3773\sqrt{9.3284}$$

$$t = -0.3773 * 3.0542$$

$$t = -1.1524$$

d: 303

n: 11



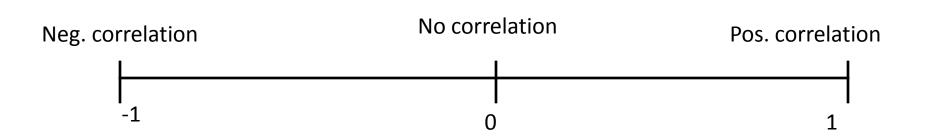
- What exactly do the number "r" and "t" mean?
 - r is the strength of correlation
 - r is always between -1 and 1

d: 303

n: 11

r: -0.3773

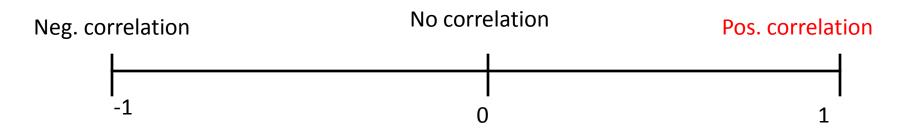
t: -1.1524





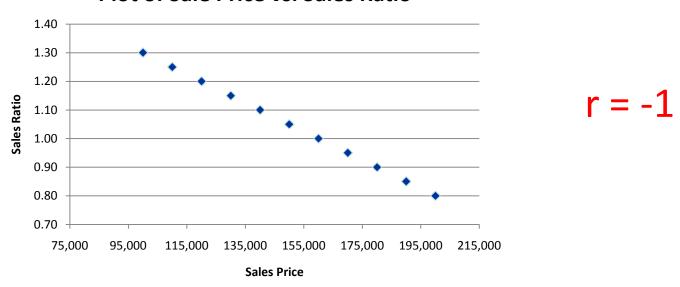
Plot of Sale Price vs. Sales Ratio

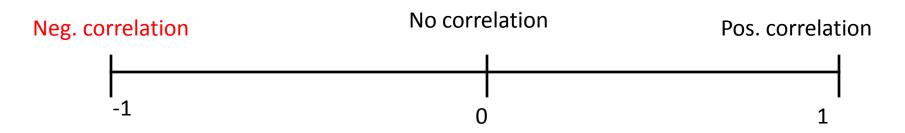






Plot of Sale Price vs. Sales Ratio

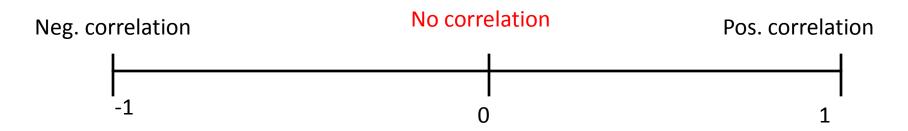






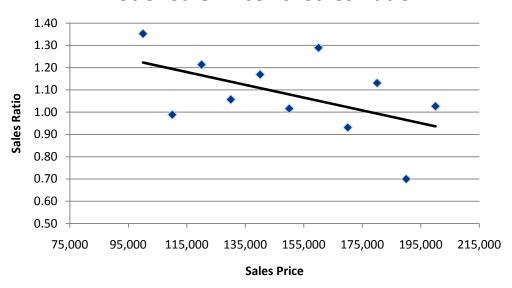
Plot of Sale Price vs. Sales Ratio



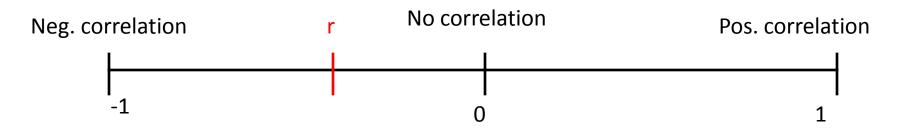




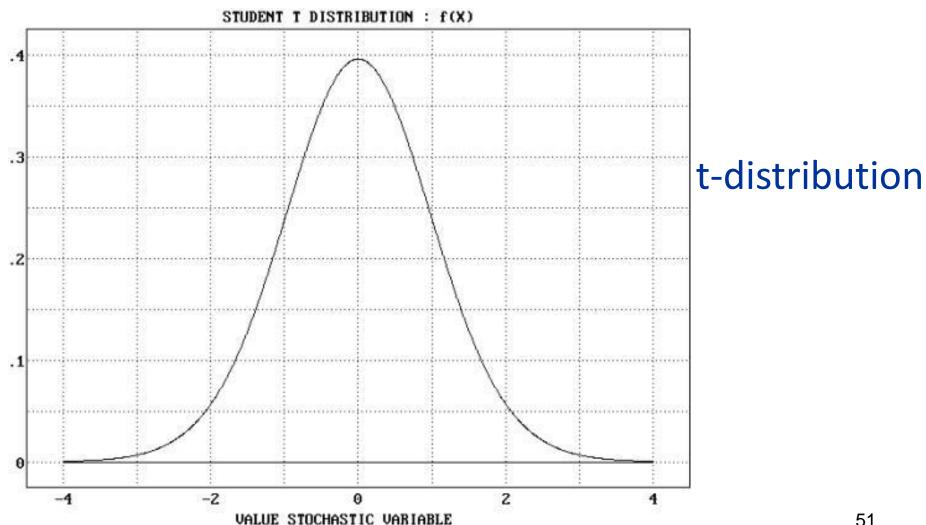
Plot of Sale Price vs. Sales Ratio



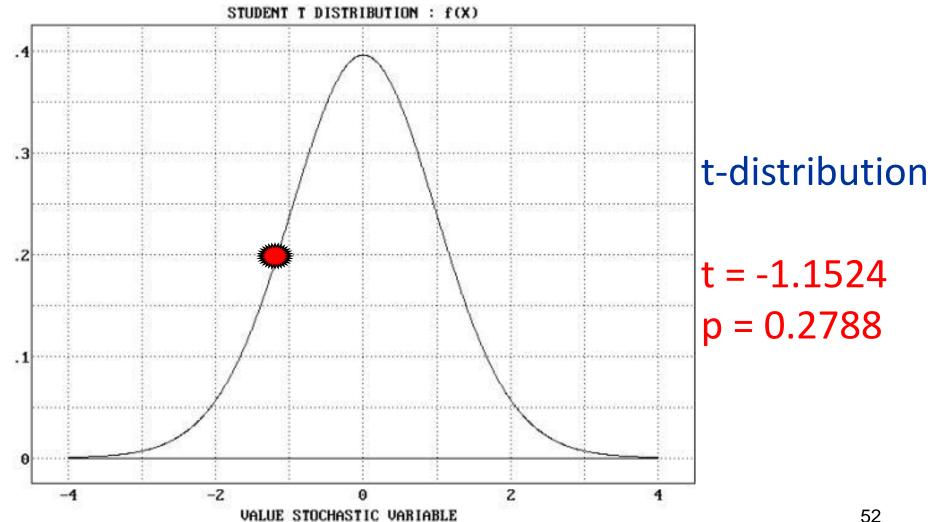
r = -0.3773













- You have assessed a group of six properties, and all statistics are fine except for the PRD, which is low. You have re-checked your data and believe the assessments are correct.
- How can you be sure there is no inequity?



Assessed Value	Sale Price	<u>Ratio</u>	Median: 1.04
35,000	56,000	0.63	COD: 16.5% PRD: 0.96
34,100	33,200	1.03	
91,000	86,400	1.05	
119,800	129,800	0.92	
136,600	101,200	1.35	
174,500	145,300	1.20	



Find the difference between the ranks.

Sale Price	<u>Rank</u>	<u>Ratio</u>	<u>Rank</u>	<u>Diff</u>	Sale Price	Ratio
56,000	6	1.03	4	2	56,000	0.63
33,200	5	0.63	6	-1	33,200	1.03
86,400	4	1.05	3	1	86,400	1.05
129,800	3	0.92	5	-2	129,800	0.92
101,200	2	1.35	1	1	101,200	1.35
145,300	1	1.20	2	-1	145,300	1.20



Square each of these differences.

					<u>Diff</u>		
Sale Price	Rank	Ratio	Rank	Diff	Squared	Sale Price	<u>Ratio</u>
56,000	6	1.03	4	2	4	56,000	0.63
33,200	5	0.63	6	-1	1	33,200	1.03
86,400	4	1.05	3	1	1	86,400	1.05
129,800	3	0.92	5	-2	4	129,800	0.92
101,200	2	1.35	1	1	1	101,200	1.35
145,300	1	1.20	2	-1	1	145,300	1.20



Add them all up.

					<u>Diff</u>		
Sale Price	Rank	Ratio	Rank	Diff	Squared	Sale Price	<u>Ratio</u>
56,000	6	1.03	4	2	4	56,000	0.63
33,200	5	0.63	6	-1	1	33,200	1.03
86,400	4	1.05	3	1	1	86,400	1.05
129,800	3	0.92	5	-2	4	129,800	0.92
101,200	2	1.35	1	1	1	101,200	1.35
145,300	1	1.20	2	-1	1	145,300	1.20
				_		_	



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6\frac{d}{n(n^2 - 1)}$$

d: 20

n: 6



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6\frac{20}{n(n^2 - 1)}$$

d: 20

n: 6



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6\frac{20}{6(6^2 - 1)}$$

d: 20

n: 6



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6\frac{20}{210}$$

$$6(6^2-1)$$

 $6(36-1)$

$$6(36-1)$$



Plug the values for "d" and "n" into this formula.

$$r = 1 - 6(.0952)$$

$$r = 1 - 0.5714$$

$$r = 0.4286$$

d: 20 n: 6



Now, plug the values for "r" and "n" into this formula.

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

d: 20

n: 6



Now, plug the values for "r" and "n" into this formula.

d: 20

n: 6

$$t = 0.4286 \sqrt{\frac{6-2}{1-(0.4286)^2}}$$



Now, plug the values for "r" and "n" into this formula.

d: 20

n: 6

$$t = 0.4286 \sqrt{\frac{4}{1 - (0.4286)^2}}$$



Now, plug the values for "r" and "n" into this formula.

 $t = 0.4286\sqrt{\frac{4}{0.8163}}$

d: 20 n: 6

r: 0.4286

 $1-(0.4286)^2$

1 - 0.1837

0.8163



Now, plug the values for "r" and "n" into this formula.

$$t = 0.4268\sqrt{4.9}$$

$$t = 0.4268 * 2.2136$$

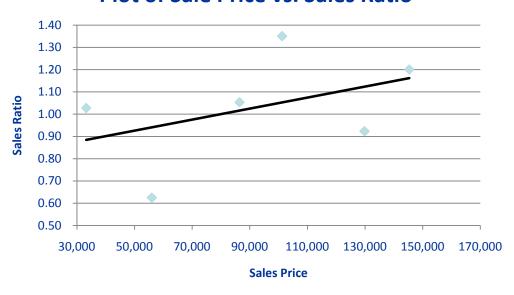
$$t = 0.9448$$

d: 20

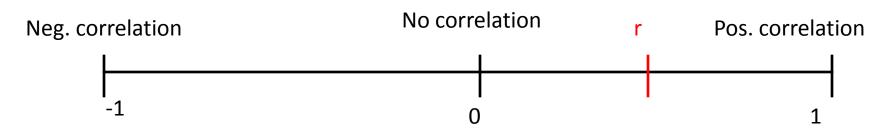
n: 6



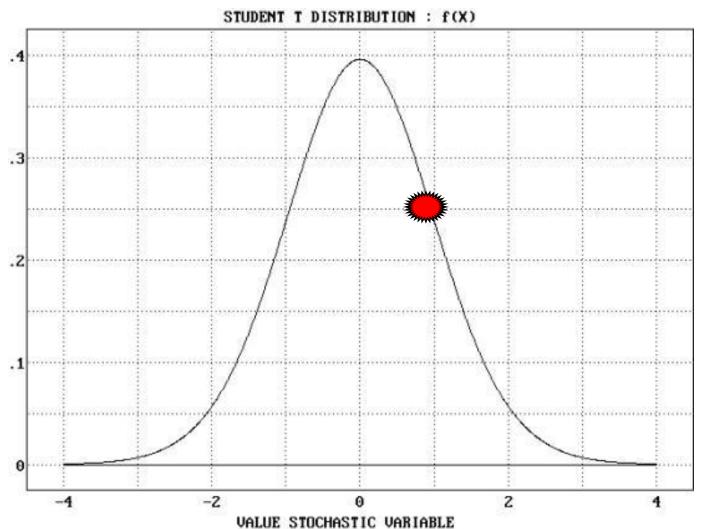
Plot of Sale Price vs. Sales Ratio



r = 0.4286







t-distribution:

t = 0.9448

p = 0.3965



Questions?



Contact the Department

- David Schwab
 - Telephone: 317.234.5861
 - Fax: 317.232.8779
 - E-mail: <u>dschwab@dlgf.in.gov</u>
- Web site: www.in.gov/dlgf
 - "Contact Us": www.in.gov/dlgf/2338.htm.